

# How To Generate a Test Fixture

The EDH Baldur project has a Fixture Maker job that generates deidentified sample of the original data that has been provided by the client.

## EDH fixture maker command examples:

For local: `edh utilities createfixture -e local -p small -n chomp-sample -r 10 -i fixture/cleanse/utilization/chomp/`

For staging: `edh utilities createfixture -e stage -p small -n chomp-large-sample -r 1000 -i media/input/utilization/chomp/`

For production: `edh utilities createfixture -e prod -p small -n chomp-qa-fixture -r 10000 -i media/input/utilization/chomp/`

## For this job we need to provide the following parameters:

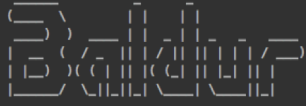
1. `-e --env` which defines the environment that the command will be executed for (local | qa | stage | prod)
2. `-p --profile` which defines the amount of capacity that will be used to execute this job (small | med | large)
3. `-n --artifact-name` refers to the name of the output fixture artifact
4. `-r --sample-rows` is the output's sample size
5. `-i --input-directory` specifies the directory containing files that needs to be deidentified and sampled

## Step 1: Run the EDH create fixture command:

```
edh utilities createfixture -e local -p small -n chomp-small-sample -r 1000 -i /Users/pallavi.karan/Documents/chomp/ -c chomp
```

## Step 2: Let the job run and finish to completion

```
Using input file '/Users/pallavi.karan/Documents/chomp/'
Using 'local' environment shell defaults from '/Users/pallavi.karan/Documents/edh-baldur/bin/defaults/local.sh'
Using 'local' environment configuration from '/Users/pallavi.karan/Documents/edh-baldur/bin/defaults/local.conf'
Job profile set to small (8 cores, 4g per node)
Executing using FULL assembly jar: '/Users/pallavi.karan/Documents/edh-baldur/target/scala-2.11/baldur-assembly-0.16.0-SNAPSHOT.jar'
```



```
v0.16.0-SNAPSHOT (Development Mode)
...*THE* Enterprise Data Pipeline...
```


```
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/Users/pallavi.karan/Documents/edh-baldur/target/scala-2.11/baldur-assembly-0.16.0-SNAPSHOT.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/Users/pallavi.karan/spark-2.1.1-bin-hadoop2.7/jars/slf4j-log4j12-1.7.16.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2018-01-04 15:50:31 INFO CreateFixtureApps:96 - Saved Dx file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/dx.csv
2018-01-04 15:51:16 INFO CreateFixtureApps:96 - Saved Facility file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/facility.csv
2018-01-04 15:52:04 INFO CreateFixtureApps:96 - Saved Visit file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/visit.csv
2018-01-04 15:52:52 INFO CreateFixtureApps:96 - Saved Cpt file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/cpt.csv
2018-01-04 15:53:40 INFO CreateFixtureApps:96 - Saved Demog file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/demog.csv
2018-01-04 15:54:22 INFO CreateFixtureApps:96 - Saved Px file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/px.csv
2018-01-04 15:55:09 INFO CreateFixtureApps:96 - Saved Guarantor file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/guarantor.csv
2018-01-04 15:56:12 INFO CreateFixtureApps:96 - Saved Physician file to location: s3a://edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20:43:36.795Z/physician.csv
wc: /Users/pallavi.karan/Documents/chomp/: read: Is a directory
```

## Step 3: Log into the staging aws console

<https://influence-non-prod.signin.aws.amazon.com/console>

Secure | <https://us-east-1.signin.aws.amazon.com/oauth?SignatureVersion=4&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAJMOATPLHVSJ563XQ&X-Amz-Date=2018-01-04T21%3A00%3A51.96...>

Apps | EDH Dev | General IH | Spark | Scala | CLI SO | Travel | Other | DataBase | Health | Prep | Github | AWS



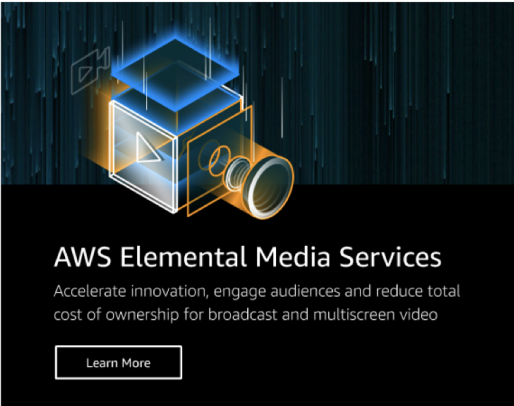
Account ID or alias

IAM user name

Password

Sign In

[Sign-in using root account credentials](#)



**AWS Elemental Media Services**


Accelerate innovation, engage audiences and reduce total cost of ownership for broadcast and multiscreen video

[Learn More](#)

English

[Terms of Use](#) [Privacy Policy](#) © 1996-2018, Amazon Web Services, Inc. or its affiliates.

#### Step 4: Go to S3 Service in aws console



Services ▾ Resource Groups ▾

### AWS services

**S3**

Scalable Storage in the Cloud

**Athena**


Query Data in S3 using SQL

**Snowball**

Large Scale Data Transport


**Amazon Transcribe**

Powerful Speech Recognition




**Launch a virtual machine**

With EC2 or Lightsail  
~1-2 minutes




**Build a web app**

With Elastic Beanstalk  
~6 minutes




**Host a static website**

With S3, CloudFront, Route 53  
~5 minutes




**Connect an IoT device**

With AWS IoT  
~5 minutes



**Start a development project**

With CodeStar  
~5 minutes



**Register a domain**

With Route 53  
~3 minutes

[See more](#)

#### Step 5: Click on the *edh-deidentified-fixture-staging* bucket

<input type="text"/> Search for buckets			
<a href="#">+ Create bucket</a> <a href="#">Delete bucket</a> <a href="#">Empty bucket</a>	40 Buckets 0 Public 4 Regions		
ai-redis-backup-staging	Error	US East (N. Virginia)	Oct 1, 2017 2:43:10 PM GMT-0400
ai-staging-es-snapshot	Error	US East (N. Virginia)	Sep 8, 2017 5:00:20 PM GMT-0400
ai-staging-es-snapshot-logs	Error	US East (N. Virginia)	Dec 26, 2017 2:03:13 PM GMT-0500
ai-staging-mongo-backup	Error	US East (N. Virginia)	Sep 11, 2017 4:39:02 PM GMT-0400
ansible-predict-backups	Error	US East (N. Virginia)	Sep 12, 2017 3:53:24 AM GMT-0400
aws-glue-scripts-283399209552-us-east-1	Error	US East (N. Virginia)	Nov 13, 2017 1:30:26 PM GMT-0500
aws-logs-283399209552-us-east-1	Error	US East (N. Virginia)	Jul 20, 2017 5:22:25 AM GMT-0400
cf-templates-elnh24h1pvn-us-east-1	Error	US East (N. Virginia)	Jul 18, 2017 7:53:40 AM GMT-0400
cf-templates-elnh24h1pvn-us-east-2	Error	US East (Ohio)	Aug 9, 2017 4:45:24 AM GMT-0400
cf-templates-elnh24h1pvn-us-west-1	Error	US West (N. California)	Aug 4, 2017 1:49:05 AM GMT-0400
cf-templates-elnh24h1pvn-us-west-2	Error	US West (Oregon)	Jul 19, 2017 4:38:48 AM GMT-0400
cloudtrail-283399209552-oregon	Error	US West (Oregon)	Jul 20, 2017 6:23:50 AM GMT-0400
edh-data-staging	Error	US East (N. Virginia)	Oct 30, 2017 4:35:32 PM GMT-0400
edh-deidentified-fixtures-staging	Not public *	US East (N. Virginia)	Nov 6, 2017 3:30:29 PM GMT-0500

Step 6: Locate the name of the artifact used while running the edh job appended by the timestamp it was created at

Services
Resource Groups

pallavi.karan @ Influence-non-...
Global
Support

Amazon S3
>
edh-deidentified-fixtures-staging

Overview
Properties
Permissions
Management

Type a prefix and press Enter to search. Press ESC to clear.

[Upload](#)
[+ Create folder](#)
[More](#)

US East (N. Virginia)

Viewing 1 to 11

<input type="checkbox"/>	Name ↑	Last modified ↑	Size ↑	Storage class ↑
<input type="checkbox"/>	chomp-small-sample_2018-01-04T20:43:36.795Z	--	--	--

Step 7: Find the de-identified & sampled files to download individually by clicking on each file and clicking the download/open option OR right click on each file and select download as which will show a pop-up and a link to download file, click on the link and it will download the file.

Type a prefix and press Enter to search. Press ESC to clear.

Upload

Create folder

More

US East (N. Virginia)

Viewing 1 to 9

<input type="checkbox"/>	Name	Last modified	Size	Storage class
<input type="checkbox"/>	temp	--	--	--
<input type="checkbox"/>	cpt.csv	Jan 4, 2018 3:52:52 PM GMT-0500	28.6 KB	Standard
<input type="checkbox"/>	demog.csv	Jan 4, 2018 3:53:40 PM GMT-0500	212.0 KB	Standard
<input type="checkbox"/>	dx.csv	Jan 4, 2018 3:50:32 PM GMT-0500	27.4 KB	Standard
<input type="checkbox"/>	facility.csv	Jan 4, 2018 3:51:17 PM GMT-0500	74.3 KB	Standard
<input type="checkbox"/>	guarantor.csv	Jan 4, 2018 3:55:10 PM GMT-0500	130.4 KB	Standard
<input type="checkbox"/>	physician.csv	Jan 4, 2018 3:56:12 PM GMT-0500	9.5 KB	Standard
<input type="checkbox"/>	px.csv	Jan 4, 2018 3:54:23 PM GMT-0500	27.5 KB	Standard
<input type="checkbox"/>	visit.csv	Jan 4, 2018 3:52:04 PM GMT-0500	130.0 KB	Standard

Viewing 1 to 9

aws

Services

Resource Groups

★

🔔

pallavi.karan @ Influence-non-...

Global

Support

Amazon S3 > edh-deidentified-fixtures-staging / chomp-small-sample\_2018-01-04T20:43:36.795Z

Overview

🔍 Type a prefix and press Enter to search. Press ESC to clear.

📁 Upload

➕ Create folder

More

<input type="checkbox"/>	Name ↕	Last modified ↕	Size ↕
<input type="checkbox"/>	temp	--	--
<input type="checkbox"/>	cpt.csv	Jan 4, 2018 3:52:52 PM GMT-0500	28.6 KB
<input type="checkbox"/>	demog.csv	Jan 4, 2018 3:53:40 PM GMT-0500	212.0 KB
<input type="checkbox"/>	dx.csv	Jan 4, 2018 3:50:32 PM GMT-0500	27.4 KB
<input checked="" type="checkbox"/>	facility.csv	Jan 4, 2018 3:51:17 PM GMT-0500	74.3 KB
<input type="checkbox"/>	guarantor.csv	Jan 4, 2018 3:55:10 PM GMT-0500	130.4 KB
<input type="checkbox"/>	physician.csv	Jan 4, 2018 3:56:12 PM GMT-0500	9.5 KB
<input type="checkbox"/>	px.csv	Jan 4, 2018 3:54:23 PM GMT-0500	27.5 KB
<input type="checkbox"/>	visit.csv	Jan 4, 2018 3:52:04 PM GMT-0500	130.0 KB

Download as

Right-click the download link below and choose "Save link as..."

facility.csv

Amazon S3 > edh-deidentified-fixtures-staging / chomp-small-sample\_2018-01-04T20:43:36.795Z

cpt.csv Latest version ▾

Overview

Properties

Permissions

Open

Download

Download as

Make public

Copy path

Owner

aws-non-prod

Last modified

Jan 4, 2018 3:52:52 PM GMT-0500

Etag

a1e2f6002d958ead386802cb6491db51

Storage class

Standard

Server side encryption

AES-256

Size

29292

Link

[https://s3.amazonaws.com/edh-deidentified-fixtures-staging/chomp-small-sample\\_2018-01-04T20%3A43%3A36.795Z/cpt.csv](https://s3.amazonaws.com/edh-deidentified-fixtures-staging/chomp-small-sample_2018-01-04T20%3A43%3A36.795Z/cpt.csv)